HAL open science

Social Engineering and Organisational Dependencies in Phishing Attacks

Ronnie Taib, Kun Yu, Shlomo Berkovsky, Mark Wiggins, Piers Bayl-Smith

To cite this version:

Ronnie Taib, Kun Yu, Shlomo Berkovsky, Mark Wiggins, Piers Bayl-Smith. Social Engineering and Organisational Dependencies in Phishing Attacks. 17th IFIP Conference on Human-Computer Interaction (INTERACT), Sep 2019, Paphos, Cyprus. pp.564-584, 10.1007/978-3-030-29381-9_35 . hal-02544575

HAL Id: hal-02544575

https://inria.hal.science/hal-02544575v1

Submitted on 16 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Social Engineering and Organisational Dependencies in Phishing Attacks

Ronnie Taib[1[0000-0001-7535-5875]], Kun Yu[2[0000-0001-5138-6749]], Shlomo Berkovsky[3[0000-0003-2638-4121]], Mark Wiggins[3[0000-0002-6422-9475]] and Piers Bayl-Smith[3[0000-0001-8014-0633]]

[1] Data61 – CSIRO, Eveleigh, Sydney, Australia
[2] University of Technology Sydney, Australia
[3] Macquarie University, Sydney, Australia

ronnie.taib@csiro.au, kun.yu@uts.edu.au,
{shlomo.berkovsky, mark.wiggins , piers.bayl-smith}@mq.edu.au

Abstract. Phishing emails are a widespread cybersecurity attack method. Their breadth and depth have been on the rise as they target individuals and organisations with increased sophistication. In particular, social engineering in phishing focuses on human vulnerabilities by exploiting established psychological and behavioural cues to increase the credibility of phishing emails. This work presents the results of a 56,000-participant phishing attack simulation carried out within a multi-national financial organisation. The overarching hypothesis was that strong cultural and contextual factors impact employee vulnerability. Thus, five phishing emails were crafted, based on three of Cialdini's persuasion principles used in isolation and in combination. Our results showed that Social proof was the most effective attack vector, followed by Authority and Scarcity. Furthermore, we examined these results in the light of a set of demographic and organisational features. Finally, both click-through rates and reporting rates were examined, to provide rich insights to developers of cybersecurity educational solutions.

Keywords: Cybersecurity, Phishing, Social engineering, Simulation, Behavioural study.

1 Introduction

In 2017, the average global annualised cost of cybercrime was $11.7M per organisation, representing a 22.7% increase over the previous year [28]. Phishing and social engineering were identified as the second highest type of attack after malwares. Phishing is a scalable act of deception whereby impersonation is used to obtain information from a target individual [15]. In practical terms, this involves sending malicious emails with the intent to deceive recipients and lure them into disclosing personal information or revealing their credentials. To increase credibility, phishing emails often mimic the design and content of genuine emails sent by reputable companies, government agencies, or personal contacts of the recipient.

The breadth and depth of phishing attacks has expanded continuously in both individual and organisational settings. For the individual, the most destructive

consequences will be identity theft or financial loss through electronic banking services. However, a successful attack on a company or an organisation can let the attacker access precious enterprise data and potentially cause significant financial, reputational, and security damage [28]. Hence, governments and financial institutions are seen as lucrative targets for phishing.

To mitigate potential losses, employees are continuously being trained to identify and report phishing attacks. But, in this 'armament race', phishing attacks have also been evolving, deploying more complex technical and psychological approaches. While early phishing methods primarily capitalised on mass-mailing of dubious content, recent methods involve the tailoring of emails to the recipients (spear-phishing or whaling), spoofing (links and email address manipulation), and even large-scale methods, such as website forgery [18]. The attack vectors have also broadened significantly, with SMS and phone phishing being quite common.

To craft a phishing attack, one of the method most widely utilised to manipulate people into responding involves social engineering [14]. Social engineering exploits established psychological and behavioural principles to increase the credibility of emails and lessen the motivation to engage in the careful evaluation of incoming information, thereby increasing the likelihood of a successful attack. Among the methods widely-encountered in phishing emails are established persuasive principles such as authority or scarcity, which are exploited for the purposes of deception rather than persuasion [17].

As phishing poses a threat to organisations and individual users, much effort has been devoted to developing tools that protect the recipients of phishing emails and help them mitigate phishing attacks. Most of these are automated tools that scan and filter incoming emails for suspicious content (keywords, grammatical mistakes), or technical issues (inconsistent links, spoofed metadata, illegitimate senders) typical for phishing attacks. While such filters achieve substantially high levels of accuracy, they also have limitations, leaving users vulnerable to targeted or not yet detected types of attacks. Therefore, there is a need to empower humans to detect malicious emails, helping them to mitigate the risks associated with phishing.

Appropriate responses to phishing emails can be achieved through education. It has been demonstrated that less technically knowledgeable users are more vulnerable to phishing attacks [22]. We posit that customised content and regime can improve current education delivery methods. To achieve such customisation at a large scale, it is crucial to first identify solid predictors of vulnerability. Past research has touched upon basic demographic factors such as age and gender, but with mixed results, potentially due to small size and methodological limitations [21].

In this paper, we present and analyse the results of a large-scale phishing attack simulation carried out by a multi-national financial organisation, as a follow-up of a qualitative study [7]. The simulation involved virtually all the employees of the organisation (more than 56,000 staff) and five variants of a phishing attack, each deploying a different social engineering strategy or a combination of strategies. The simulated phishing attack was initiated by the organisation in a drill-style exercise, and emails were sent directly to the employees' corporate email addresses. We captured the employee responses to the phishing emails: *reporting* the attack, *clicking* on the link in the phishing email, or *ignoring* the email.

We report the observed responses of the employees and analyse them with respect to multiple demographic features, organisational categories, and the deployed social engineering strategies. To the best of our knowledge, this is among the first works to look beyond demographic factors into the more specific employment and organisational factors at such a large-scale, diverse role- and age-wise, and multi-national corporate environment. The outcomes surface important insights that should be considered by developers of future, tailored phishing education programs.

In summary, the contributions of our work are three-fold. First, we highlight demographic segments and categories of employees that are vulnerable to phishing attacks. Second, we compare the deceptive power of several social engineering strategies deployed for phishing purposes. Third, we discuss how organisations can operationalise these findings to develop tailored education campaigns and help employees mitigate phishing attacks.

2 Related Work

Phishing is an activity covering a range of media, approaches, attack vectors, and objectives. Hence, it has been difficult to establish a common definition for phishing, although Lastdrager, based on the lexical and semantic analysis of the literature, coined the following: "*phishing is a scalable act of deception whereby impersonation is used to obtain information from a target*" [15]. In this paper, we focus on phishing emails in a context where the information obtained from the target is for a malicious intent, such as to gain financial benefit or compromise an organisation's data or reputation.

Phishing emails typically contain a link to a fraudulent website or include malicious attachments. They often utilise a range of techniques to reach their target and entice them to disclose information. Protective methods can broadly be categorised into technological (filtering emails based on spoofed sender address or known harmful content), organisational (policies and procedures for recruitment, technology use) and human (training and educating users) [12,18]. While all three categories must be addressed holistically to protect an organisation [12], this paper centres on human factors and social engineering attack vectors.

Social engineering refers to deception methods leveraging psychological mechanisms that reduce suspicion and increase trust in the malicious content. Several studies have attempted to define a taxonomy of social engineering attacks. At the highest level, they can be deconstructed into type (e.g., using social attributes to persuade the targets to divulge information), channel (e.g., email) and operator (e.g., sent by a purported human), and applied through a specific attack vector such as phishing [14].

Seamlessly persuading a target to disclose information is the main pillar of phishing attacks. Hence, the six basic principles of influence, originally proposed by Cialdini [6] in the field of marketing and human persuasion, can also be applied for phishing as "misused weapons" [1,17]. The six core principles of persuasion are:

- Reciprocation: people tend to feel indebted to someone who did them a favour, e.g., they will positively regard a notification of account expiry and may enter their credentials to correct it;

- Commitment and consistency: people tend to honour their commitments, e.g., reviewing the bill of hotel they previously booked;
- Social proof: people tend to trust a source if others are believed to also trust that source;
- Liking: people tend to trust more a person they know or some content that looks familiar to them;
- Authority: people tend to comply with authority, be it in the form of a person's position, or an organisation's logo.
- Scarcity: people tend to hesitate less and act sooner if they believe that there is a limit in amount or time to obtain something of interest.

Authority is identified as one of the most effective principles in a number of studies [1,3,5,11,19], while the other principles may be more influential, depending on other characteristics. For example, while scarcity has been demonstrated to influence younger people, reciprocation is likely to influence older people [19]. Furthermore, the use of these principles by attackers is evolving over time, with an increasing use of scarcity, whereas reciprocation and social proof have been declining in usage [27].

Other efforts were made to combine Cialdini's principles with psychological profiling, e.g., the Big Five model [2], or with Gragg's psychological triggers and Stajano et al.'s principles of scams, leading to the principles of persuasion in social engineering [11]. The latter work lists and analytically compares the principles outlined by Cialdini, Gragg, and Stajano et al., showing that the three sets overlap for the authority and social proof principles, where distraction (scarcity) and authority were the most effective principles [10]. Combining the results from these studies, we selected social proof, scarcity and authority as the principles to deploy in the present study.

Considering the 'scalable' part of the phishing definition, we note that most of the existing studies are limited to a few hundred participants, often university students [4,5,8,9,19,20,23,25]. A few studies included larger samples, such as Jagatic et al. with 1,731 participants in total (although only about 600 participated in the phishing component of their study) [13]. Mohebzada et al. engaged more than 10,000 students, staff and alumni [16]; however, again in a university environment. Sheng et al. collected 1,001 user responses, but noted that the use of crowdsourcing might have impacted the diversity of their sample population.

The only study using a large corporate-type of population has been reported by Williams et al. [26]. They used the results of nine phishing simulations administered to 62,000 public service employees, although they did not have access to the demographic characteristics of their participants. They focused on urgency and authority in isolation and showed that their application increased the effectiveness of spear phishing attacks. Our study extends [26] by also focusing on Social proof, as we posit that, in organisations, with a strong employee-identification culture, users may be at greater risk of Social proof, rather than of other types of attacks. Moreover, we examine the correlations between persuasion principles, demographic characteristics, and organisational parameters of the employees to test the links between multiple factors affecting vulnerability to phishing attacks.

Overall, while there has been some consideration of demographics in previous research, the results are inconclusive. In the context of gender, some studies report females as more vulnerable than males [8,13,19,22], while others report the opposite [16]. In the context of age, younger participants appear most at risk [8,13,22], at least

when exposed to scarcity-based material [19]. However, contradictory results are also reported, whereby young participants are less likely to fall victim [16]. To the best of our knowledge, only one study examined the links to corporate variables, namely tenure [4]. As expected, tenure positively correlated with age, but potentially, was a stronger predictor of phishing vulnerability. Correlations were negative, such that on average, more years of service were associated with a decrease in vulnerability.

Two main methodologies were used in previous research. Role-play studies typically present images of phishing emails to users and seek respondents' intended actions, e.g., clicking the link or deleting the email [9,22,23]. By contrast, drill-style studies send actual phish emails to participants and monitor their responses [4,9,13,19,25]. In some cases, participants were fully aware of the study taking place, e.g., because they installed specific monitoring software on their machine [19], while, in other cases, they were completely unaware of being used as participants [16], which caused post-study discussions [13]. To the best of our knowledge, the ecological validity of the role-play methods has not been fully investigated in prior works. Hence, drill-style studies with minimal warning seem to offer the highest levels of validity. Such warnings can materialise as a general email sent to an organisation, yet keeping in mind that a significant proportion of recipients may simply ignore the email [16].

3 Objectives

Building on existing research, this paper is part of a holistic approach to cybersecurity, aimed at empowering end-users to detect and respond to phishing attacks. We posit that the one-size-fits-all education approach, commonly used in large organisations misses its objective because it does not address the various needs and attitudes of employees. Hence, the overarching motivation of our work is to design tailored educational material crafted with various user profiles in mind.

However, the first step in designing such tailored education approaches lays in understanding the dependencies between demographic characteristics as well as organisational roles of the employees, and their vulnerability to various types of phishing attacks. To this end, in this work we set out to validate several hypotheses:

- Different age groups fall victim to different types of persuasion, and therefore, should receive educational content focusing on these specific types;
- Vulnerability decreases with tenure as employees become more familiar with the processes and structure of the organisation, and therefore, they are less likely to be vulnerable to an attack;
- Employees with managerial responsibility are more likely to exert caution than non-managers. Hence, the motivational messages for these groups would vary, e.g. focusing on the company image for manager, or on the consequences of non-compliance for non-managers.

Validating these dependencies would allow us to quantify individual or group-based vulnerability of employees to phishing attacks. In turn, not only this allows to tune the sensitivity of cybersecurity technologies, e.g., of the phishing filter deployed in email clients, but also the educational campaigns and cybersecurity trainings can be tailored to the specific risks faced by the target group of employees.

4 Experiment Setting

We first present the experimental setting, simulated phishing attack, data collection methods, and the evaluation metrics that were used in this study.

4.1 Participants

In this work, we carried out a controlled, organisation-wide phishing simulation. The organisation at hand is a multi-national financial institution operating in more than 30 countries. The present study included response data from 20 countries (Australia, Cambodia China, Fiji, Hong-Kong, India, Indonesia, Japan, Lao, New-Zealand, Papua New Guinea, Philippines, Samoa, Singapore, Solomon Islands, Taiwan, UK, USA, Vanuatu, Viet Nam), with data from 10 countries excluded due to low numbers. The organisation regularly conducts internal simulations and education campaigns with the intention of raising employee awareness of information security breaches, cybersecurity attacks, and phishing emails. As the attack simulation reported in this paper was, in part, conducted for research purposes, the experimental design and data collection methods were reviewed and approved by an accredited national Human Research Ethics Committee independent of the research team.

The reported phishing simulation was sent to more than 56,000 recipients from all departments and business units of the organisation in October 2017 and results collected over the following two weeks. Descriptive statistics of the demographic segments and employment categories are provided in Table 1. As the social engineering strategy was chosen randomly for each participant, the distribution of strategy type is uniform. In the context of gender, there were approximately equal numbers of female and male participants. The data were binned into three similar-sized age groups. For tenure, we used industry standards of 5 and 10 years to group the employees. As can be expected in most organisations, the distribution of managers versus non-managers is not balanced. Only 15% of the participants were managers (that is, had employees reporting to them), while 85% were non-managers.

Table 1: Distribution of the participants (total N=56,365).

Feature	Distribution	Count	% population
Social engineering strategy	Social proof	11,268	20.0%
	Scarcity	11,267	20.0%
	Authority	11,285	20.0%
	Social proof + scarcity	11,281	20.0%
	Social proof + authority	11,264	20.0%
Gender	Female	28,430	50.4%
	Male	27,935	49.6%
Age	18-31	19,502	34.6%
	32-40	19,354	34.3%
	41+	17,509	31.1%
Tenure	0-5 years	32,032	56.8%
	5-10 years	11,714	20.8%
	10+ years	12,619	22.4%
Manager	Managers	8,376	14.9%

Non-managers	47,989	85.1%

4.2 Methods

We deployed five variants of phishing email, all conforming with three of Cialdini's core principles of persuasion – *social proof*, *scarcity*, and *authority* – or their combinations. As explained earlier, these principles were selected based on the combination of prior works. Moreover, it can be noted that the other three principles – reciprocity, consistency, and liking – were considered less appropriate in the context of our simulation because (i) the one-off nature of phishing emails naturally could not leverage existing relationships and social ties between the sender and the recipient; and (ii) the risk of compromising the reputation of real employees or managers precluded us from using real names or roles within the organisation. As Cialdini's principles of persuasion are used in phishing attacks to deceive email recipients rather than to persuade them, we will refer to these hereafter as *social engineering strategies*.

In addition to the three core principles, we also considered two combinations of principles: *social proof* + *authority*, and *social proof* + *scarcity*. The authority + scarcity combination was considered but not studied, due to its low compatibility and less realistic outcome, when co-located in the same email. Consequently, we produced five variants of phishing emails. To verify their alignment with Cialdini's principles and with the organisation's corporate communication style, the crafted phishing emails were reviewed and iteratively revised by five human-computer interaction researchers and the security team of the organisation. The input of the latter ensured that the simulated phishing emails were broadly consistent with real phishing emails that had been previously detected at the organisation. Finally, minor typographic errors, grammatical mistakes, and language inconsistencies were intentionally introduced and uniformly distributed amongst the emails.

All of the emails were sent from an external email address with a domain name that closely resembles the organisation's, but has a minor typographic error. All the emails included a URL linking to the same domain. Although the domain name modification was minor, it was expected to be readily recognised by the employees of the organisation, frequently visiting web and intranet sites hosted by the original domain of the organisation. The emails were white-listed by the mail servers of the organisation and, as such, all reached their recipients.

Key excerpts from the five phishing emails are copied verbatim in the blocks below. For illustration purposes, we highlight those parts of the email that are associated with the deployed social engineering strategies. Linda Gardner, who is signed on the emails, is a fictitious name and not an employee of the host organisation.

<u>Social proof</u>

Check out this offer:Dinner carnival is coming! – based on recent survey, ***more than 80% of our staff like*** *to dine out with friends and family, soo we have been negotiating with high-end local restaurants to bring you a meal discounted at 95% between now and the end of next week. I hope you'll take-up this unique chance to enjoy the delicious food and lovely dining environment.*

*So far **over 100 staffs have enjoyed this discount**. Please check [LINK] to find out which of your local restaurants are available for this nice offer and regster accordingly.*

Linda Gardner

Scarcity

Check out this offer:Dinner carnival is coming! – we have been negotiating with high-end local restaurants to bring you a meal discounted at 95% between now and the end of next week. So I hope you'll take-up this unique opportunity to enjoy the delicious food and lovely dining environment.

*However, opportunties are limited - only the **first 200 registrations** will be able to secure the offer. So, hurry up and register at [LINK]*

*The **registration may end anytime in the next couple of hours** before COB today, soo be quick!*

Linda Gardner

Authority

*The Life Balance Team has announced an annual Dinner carnival. Based on the findings of **Prof. Clark** and the Life Balance team, we want to be the perfect work-life balance place. We have been negotating with high-end local restaurants to bring you a meal discounted at 95%. So, we suggest you take-up this unique chance to enjoy the delicous food and lovely dining environment.*

***We strongly recommend** you check [LINK] to find out which of your local restaurants are available for this offer and register accordingly. **This is an organisational priority for us**, so we expect you to take part.*

Linda Gardner

Chief Officer

Social proof + scarcity

*Dinner carnival is coming! – based on recent survey, **more than 80% of our staff** like to dine out with friends and family, so we have been negotating with high-end local restaurants to bring you a meal discounted at 95% between now and the end of next week. So, I hope you'll take-up this unique opportunity to enjoy the delicous food and lovely dining environment.*

*However, **opportunities are limited** - only the **first 200 registrations** will be able to secure the offer. So, hurry up and register at [LINK]*

Linda Gardner

Social proof + authority

*The Life Balance Team has announced an annual Dinner carnival. Based on the findings of **Prof. Clark** and the Life Balance team, we want to be the perfect work-life balance place. Based on a recent survey, **more than 80% of our staff** like to dine out with friends and family, so we have been negotating with high-end local restaurants to bring you a meal discounted at 95%. So, we suggest you take-up this unique chance to enjoy the delicous food and lovely dining environment.*

So far ***over 100 staffs have enjoyed this discount****.* ***We strongly recommend*** *you check link to find out which of your local restaurants are available for this offer and register accordingly.* ***This is an organisational priority for us****, so we expect you to take part.*

Linda Gardner

Chief Officer

4.3 Metrics

Upon receiving the emails, each participant could respond in three ways. The first would be to click on the link, virtually falling victim to the phishing attack. This is the undesired outcome of the simulation. In this case, the participant would be re-directed to an internal educational page articulating that the email was sent as part of a phishing attack simulation conducted by the organisation. Of course, clicking on a link may not result in users necessarily disclosing their credentials, but they are still vulnerable to drive-by download malware on some websites and other types of attacks which could compromise the organisation's security. Therefore, consistent with other research studies we considered clicking on a link as a risky behavioural response.

The second response, which is the desired outcome of the simulation, would be to report the email as a phishing attack. In this case, the participant would be automatically notified by email that they successfully recognised the attack and passed the phishing simulation. The host organisation considers reporting to be far superior to ignoring attacks because it can lead to swift protection of the whole organisation thanks to a snowball effect. From a psychological standpoint, it also reflects a totally different attitude: a healthy appreciation of the risk and locus of control, rather than perceiving the threat severity of phishing attacks as low.

The third response is essentially not to respond – neither report nor click – or to ignore the phishing email. It may reflect an undesirable condition where reporting is not performed because a user is 'unsure', which highlights flaws in the promotion of 'there is no dumb question when it comes to cyber'. This aspect could not be ascertained through our study data, and furthermore, the lack of reporting may not inform us as to the efficiency of social engineering attacks, so we treat it as an undesired outcome.

It should be noted that the responses of clicking and reporting are, in principle, not mutually exclusive. In fact, a small portion of participants reported phishing attacks after viewing the educational content. However, in the following analysis we discard this minor overlap.

Following this logic, we define two metrics allowing us to quantify the performance of a group of participants in response to a phishing attack. The metrics that quantify the success of the attack on a group and the success of a group in mitigating the attack are *Click Rate* (*CR*) and *Report Rate* (*RR*), respectively. More formally, let us denote by G_c and G_r the two subgroups in the entire population of phishing attack recipients G, who clicked on or reported the attack, respectively. *CR* and *RR* are computed as

$$CR = \frac{|G_c|}{|G|} \qquad RR = \frac{|G_r|}{|G|}$$

where $|\cdot|$ denotes the number of employees within a group. Essentially, these metrics quantify the portion of group members, who clicked on or reported the attack.

5 Results

In this section, we present the results of the phishing attack simulation. Consistent with the aims of the study, we analyse the differences observed with respect to a number of demographic and organisational categories, and refine these findings by considering the effectiveness of various social engineering strategies. Therefore, as appropriate when analysing significant associations between categorical variables, Chi Square statistical tests were used for all comparisons and are noted in the diagrams below as * for p<.05 and ** for p<.01. No marking indicates that the difference was not statistically significant, p>.05. When more than two groups are compared, we initially report the combined Chi Square, before presenting further analysis where each group is tested against the rest of the sample (i.e., all other groups combined).

5.1 Overall click through and reporting rates

The overall headcount of the organisation at the time of the study was 57,089 employees. After filtering errors and missing values, we obtained reliable data related to the simulated phishing attack from 56,365 participants, which accounts to 98.7% of total employees and will be considered as the overall number of participants in the analyses reported below.

In total, 6,922 participants (*CR*=12.28% of the employees), clicked on the phishing link and virtually fell victim to the attack. On the other hand, 12,219 participants (*RR*=21.68% of the employees) reported the received email as a phishing attack, thus, successfully having passed the simulation, as shown in Fig. 1. The remaining 68.87% of employees did not respond actively to the simulation. Overall, we observe that the number of reports was 76% higher than the number of clicks, generally indicating a positive preparedness of the organisation to mitigate phishing attacks.

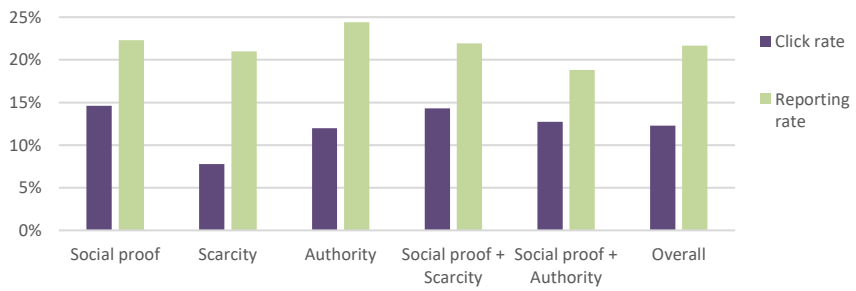

Fig. 1: Overall click-through and reporting rates.

Comparing the effectiveness of the three core social engineering strategies – Social proof, Scarcity, and Authority – we observe that the click-through rates differ substantially across these strategies, with Scarcity being perceived least credible (*CR*=7.8%), followed by Authority (*CR*=12.0%) and Social proof (*CR*=14.6%). Hence, the latter is 87.2% and 21.7% higher than the *CR*s of Scarcity and Authority, respectively. Despite the different click-through rates, the obtained reporting rates *RR* of the three core strategies are comparable, all placed within the 21.0-24.4% range.

We posit that this might be explained by the higher popularity of phishing emails leveraging Authority and Scarcity [26]. Therefore, the email recipients are familiar with this types of attack and are more likely to recognise them, as reflected by the lower *CR* of these strategies. In essence, a combination of the *CR* and *RR* results shows that *Social proof is the most effective social engineering strategy*, followed by Authority, and then by Scarcity.

Combining Social proof with the other social engineering strategies, does not provide an additive effect over and above a single strategy. The addition of either Scarcity or Authority to the Social proof strategy slightly compromises the credibility of the phishing emails, such that their respective *CR*s drop by 13.0% and 2.3% in comparison to the *CR* of Social proof deployed in isolation. Similarly, the *RR* decreases slightly when combining social engineering strategies (1.7% and 15.6%) in comparison to the *RR* of Social proof in isolation. In summary, we observe that the *addition of recognisable strategies decreases the effectiveness* of the phishing emails.

5.2 Gender

In contrast to previous studies [7,13,19,22], the analysis of gender differences (Fig. 2) shows that there were no significant differences between female and male participants' click-through rates. Our population sample is much larger and more diverse in terms of age, occupation and country of origin than in prior research, so our result may have higher ecological validity within the general population. In terms of reporting phishing emails (Fig. 3), males were significantly more likely to report phishing emails than females, except for the Social proof + Scarcity attack strategy.

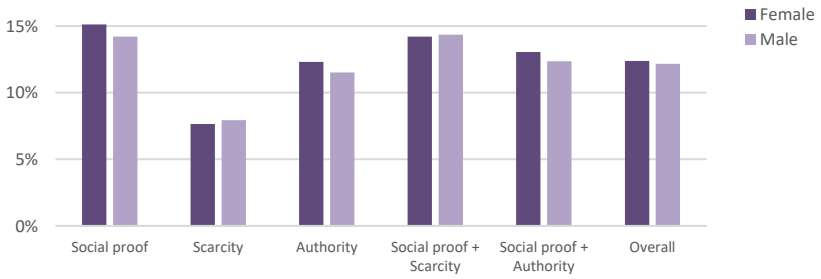

Fig. 2: Click-through rates per gender.

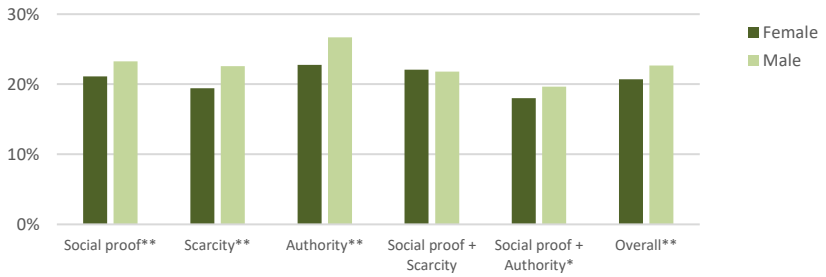

Fig. 3: Reporting rates per gender.

5.3 Age

Considering the differences observed with respect to the participants' age (Fig. 4), significant differences were evident between groups overall for all strategies except Scarcity and Authority.

Further analysis of each group versus the rest of the population provides more insights: the 41+ group is significantly more likely to fall victim to Social proof attacks than the other employees ($p<.01$). For Social proof + Scarcity, all age groups are significantly different from the rest of the employees ($p<.05$). Both the 18-31 and 41+ groups are significantly less likely to click than the rest of the population under the Social proof + Authority attack, but this is mostly due to an abnormally high click-through rate for the middle group.

In the context of reporting phishing emails (Fig. 5), we observe significant differences ($p<.01$) between the different age groups for Social proof, Scarcity, and overall. Further analyses show that the 18-31 group is significantly more likely to report Social proof ($p<.01$) or Scarcity ($p<.05$) attacks, while the group of 41+ employees are significantly less likely to report phishing using the same strategies (all $p<.01$, except Scarcity for 18-31: $p<.05$).

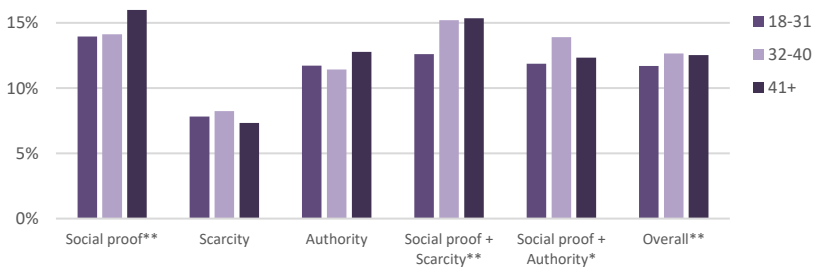

Fig. 4: Click-through rates per age.

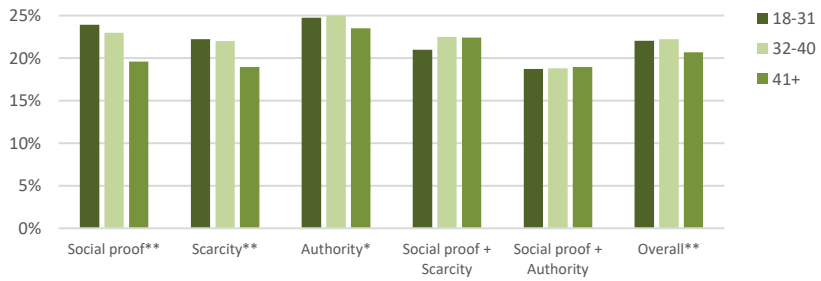

Fig. 5: Reporting rates per age.

These results indicate that *older participants are more vulnerable to phishing* than others. We posit that this observation might be attributed to lower a familiarity of the older employees with IT practices within the organisation and changes in modern technologies in general. This outcome is consistent with previous research examining phishing vulnerability of older adults [14].

5.4 Tenure

We now turn to the differences observed with respect to the number of years of service (tenure). While previous demographics-based analyses are applicable to the society at large, the following ones may be more relevant to large-scale organisations with established hierarchical structures.

Employees were grouped into three categories, including tenure of up to 5 years, 5 to 10 years, and more than 10 years (Fig. 6). We observed significant differences overall ($p<.01$) for all attack strategies. Further analyses also reveal significant differences between each group and the rest of the population for all types of attacks, apart from the 5-10 years of tenure employees and the rest for Scarcity or Social proof + Authority. In other words, the up to 5 years of tenure group is significantly more likely to click through than the rest of the population.

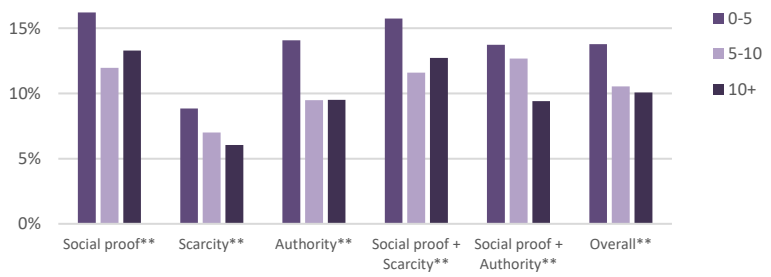

Fig. 6: Click-through rates per tenure (3 groups).

This is an interesting finding when contrasted with the age-focussed results discussed previously. Age in isolation may not be a reliable indicator of vulnerability, since organisational interventions are likely to override their effects. For example, regular training and awareness campaigns potentially decrease the vulnerability of staff during their first few years at the company, regardless of their age when joining. We tested this hypothesis further by grouping tenure into 9 comparable-size bins as shown in Fig. 7. This analysis allows us to isolate the group of new employees in their first year with the organisation. These employees are significantly more likely to click on phishing emails (overall and for each strategy: $p<.01$) than other staff.

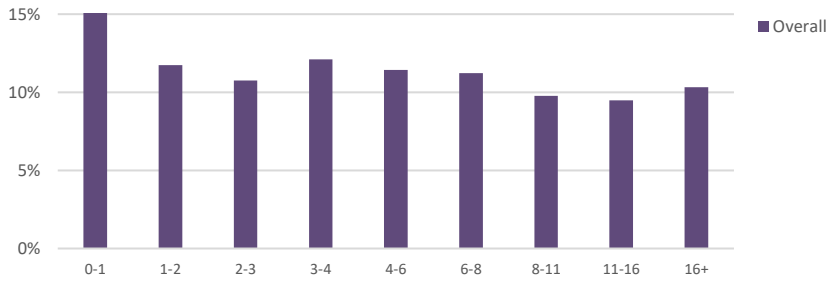

Fig. 7: Click-through rates per tenure (9 groups).

In the context of reporting rates (Fig. 8), the trend holds for newer employees whereby they were less likely to report the phishing email than the rest of the staff. However, while employees with longer tenure exhibited a fairly consistent behaviour for click-through rates, this time, the group with 5 to 10 years of tenure exhibited significantly more reporting than the rest of the population for all types of attacks ($p<.01$) except for Scarcity, where all groups were equally likely to report.

In summary, we conclude that *new employees (and, specifically, those in their first year of employment) are more vulnerable to phishing* than employees who have been with the organisation for a longer period of time. We posit that this can be explained by the lower familiarity of the former with the organisation's emails and communication styles. This may hamper their ability to distinguish between genuine emails and phishing attacks, hence, increasing their vulnerability. This is consistent with the results obtained in prior literature [16].

Fig. 8: Reporting rates per tenure (3 groups).

5.5 Managerial status

Another important organisational feature is the manager/non-manager status, communicating whether other employees report to the phishing email recipient (Fig. 9). We observed that managers were significantly less likely to click on phishing emails overall and for all types of strategies ($p<.01$, except Authority $p<.05$) with the exception of Social proof and Social proof + Authority ($p>.05$).

For reporting rates (Fig. 10), managers were significantly more likely to report phishing emails overall, and also for all types of attacks ($p<.01$, except for Scarcity and Social proof + Authority: $p<.05$).

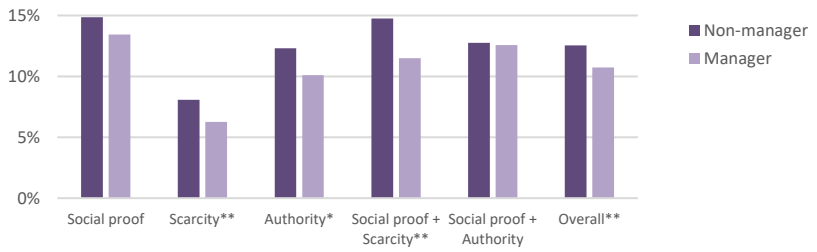

Fig. 9: Click-through rates per managerial status.

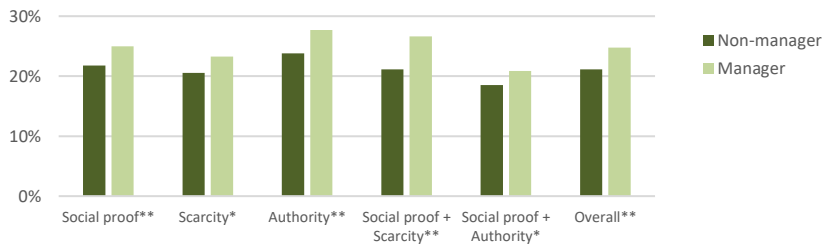

Fig. 10: Reporting rates per managerial status.

We conclude that *non-managers are more vulnerable to phishing attacks* than managers. This outcome may be explained by the higher familiarity of managers with the organisation, which allows them to better recognise and mitigate phishing attacks.

6 Discussion

We conducted a study of phishing attacks and vulnerability within a financial organisation. This study is uniquely large and diverse, with staff spanning across 20 countries and several cultures. It focuses on a number of demographic and organisational features, as well as on the social engineering strategies deployed. The main insights are:

- Females and males did not differ in their click-through rates of phishing emails, which contrasts previous research carried out with smaller or less diverse population samples. However, male employees demonstrated a higher rate of reporting phishing emails than females.
- Employees aged over 41 were generally more vulnerable to phishing attacks than other employees. They were particularly vulnerable to attacks deploying the Social proof strategy, as identified from the highest click-through rate and lowest reporting rate. Furthermore, employees younger than 32 years old were less vulnerable to all phishing attacks involving Social proof, while the employees aged between 32 and 40 were more likely to be the victimised by phishing attacks with combined strategies involving Social proof.
- Employees who have been with the organisation for fewer than five years (and, particularly, less than a year) were more vulnerable than others to phishing attacks, and also exhibited the lowest reporting rate, which was evident across all the social engineering strategies. Overall, employees who have worked in the organisation for a long period (more than 10 years) were less likely to fall victim to phishing attacks of any kind, although they may not report suspicious emails as frequently as those employees who have worked for a moderate time (5-10 years).
- Non-managers were more vulnerable to phishing attacks than managers. This finding was also observed across the board and did not depend on the social engineering strategy deployed.

- Generally, phishing attacks utilising the social proof strategy were more effective and dangerous than attacks using other strategies. This was clearly and evidently observable for all the analysed groups and organisational roles of participants.

While the first two findings related to the demographic factors reaffirm or nuance some previously reported works, the latter three, which are relevant to large organisations, are novel and original. The two findings related to tenure and manager status can be supported by the same common logic: familiarity of employees with the organisation – be it with the communication styles or the leadership team – allows them to better distinguish between genuine and fraudulent emails, and, thereby mitigate phishing attacks. We believe that this insight can be taken up by the designers of corporate training programs, particularly when on-boarding or targeting new staff.

We also observed insightful results related to the deployment of the social engineering strategies. It was evident that the effectiveness of phishing emails depends on the social engineering strategy deployed, with socially-driven manipulations being the most effective method, noticeably dominating both scarcity- and authority-driven attacks. We attribute these differences to prior familiarity of the participants and phishing email recipients in general with the deployed strategies. For example, authority-driven phishing attacks that pretend to be sent by the police, government agencies, or large companies like Apple or Microsoft, are relatively common. Likewise, numerous attacks use fraudulent scarcity-driven manipulations like “special prices for the day” or “lucky winner of a draw”. We believe that, due to their sheer popularity, participants are naturally trained to recognise and mitigate such attacks better than the more uncommon social manipulations. In contrast, mentioning in the email that other employees of the organisation accepted an offer seems to be an effective means to deceive the recipients. A potential explanation might be that employees with a strong identification with the organisation look to decrease uncertainty by conforming to what others in their group are doing (i.e., self-categorisation theory [24]), and may be particularly susceptible to social manipulations. In other terms, companies with high organisational identification of employees may be more vulnerable to social proof than other types of deception.

The diversity observed in the vulnerability of various demographic segments and organisational roles calls for tailoring of future phishing education programs. For example, junior employees need to receive more encompassing phishing education than employees at the management levels. It is also evident that different social engineering strategies need to be emphasised for younger and older employees. Our findings undermine the validity of the current, one-size-fits-all phishing education campaigns and highlight the emergent need for group tailored (or even personalised) solutions, which can cater to the specific needs and weaknesses of every employee.

Several potential limitations of our work should be raised. The first refers to the reasonably low *a priori* effectiveness of the simulated phishing attack. The study did not tailor the phishing emails to group or individual characteristics of the recipients using demographic or organisational data available, which would have amounted to spear-phishing attacks. This was done in order not to compromise any departments within the organisation and to maintain the ecological validity of the study, because detailed personal information is typically not available at large scale to outside

attackers. Hence, the level of attack effectiveness of our emails was deliberately limited.

The second limitation to be mentioned refers to the somewhat limited response rate of the simulation. As mentioned earlier, less than 35% of participants actively responded to the emails, i.e., clicked on the phishing link or reported the attack. This limitation is alleviated, however, by the large sample size of more than 56,000 participants. Therefore, we were able to obtain solid empirical evidence and analyse differences between participants with respect to social engineering strategies, as well as demographic and organisational factors.

The third potential limitation refers to the scope of the study, which involved a single phishing attack and was carried out in a single organisation operating in the financial sector. Again, we alleviate this limitation by the large sample size of the phishing attack simulation, which ensures representativeness with respect to a variety of countries, age groups, and other demographic and organisational factors. Furthermore, we should note the broad range of positions filled by the participants, including business/data analysts, financial advisers, administration staff, call centre staff, insurance specialists, and many more. Therefore, we believe that our findings are generalisable and will be also valid in other large organisations and economy sectors.

7 Conclusion

Phishing attacks have evolved over the last two decades into powerful and dangerous cybersecurity weapons. They are exploited in a multitude of cyber-attack scenarios, targeting organisations, companies, and individuals, and can be encountered in a range of domains and applications. While numerous technical solutions to phishing attacks have been developed, we argue that human users are one of the weakest links in the cybersecurity chain. Consequently, we believe in the emergent need to educate users and upgrade them into an active defence against cyber-attacks.

Large organisations rely on phishing simulations as an awareness tool and complement for a typical annual training. However, such a single point of measurement is fairly artificial and “ticking compliance boxes” does little to change users’ attitudes and behaviours towards cybersecurity. While resisting the appeal of persuasive content may be difficult, a reliance on corporate protection can be reduced by emphasising locus of control, and conversely decreasing over-confidence in other users by managing threat appraisal. Our work departs from traditional, generic, cybersecurity educational content by identifying combinations of demographic and organisational factors that are most likely to lead to improvements in performance.

Indeed, everyone is different, and the same attack may be easily mitigated by some while victimising others. This brings forward the question of individual differences in phishing vulnerability, which we set out to investigate in this work. Our overarching hypothesis is that strong culture and context factors impact employee vulnerability. To this end, we conducted a large-scale phishing attack simulation focussing on demographic and organisational differences affecting the vulnerability to phishing, also considering the social engineering strategies deployed by phishing emails. The study was carried out within a real work environment of a multi-national financial organisation and involved more than 56,000 participants.

Our results surfaced several valuable insights related to the vulnerability of different participants to phishing attacks. While previous works linked differences in phishing vulnerability to demographic factors, to the best of our knowledge, this is among the first works to identify differences with respect to organisational features in a large corporate environment. This work also examined the deceptive potential of established persuasive strategies applied for phishing attack purposes. Finally, our work considered both click-through rates and reporting rates, which paves the way for the development of the next generation of phishing awareness tools and cybersecurity educational solutions.

One immediately realisable finding that stems from our work concerns the need to tailor or personalise future cybersecurity education programs. The differences observed between the participants clearly show the strengths and weaknesses of various segments of employees. Hence, it is beneficial to focus education on the specific weaknesses of each and every employee or group of employees, to better cater to their vulnerabilities. This task is not easy to accomplish, but it has the tremendous potential to improve the efficacy of educational programs and the individual strategies for mitigating phishing attacks.

This naturally leads to future research directions, one of which targets the evaluation of such tailored phishing education programs. We expect their efficacy to be superior to the efficacy of the non-tailored, one-size-fits-all education programs that are currently deployed by many organisations. This, however, can only be validated in a follow-up user study, comparing the observed changes in the responses to phishing attacks. We will embark on this research after developing the new education program.

Another promising direction refers to the use of alternative statistical models and approaches to data analysis, such as the generalised linear model frameworks and machine learning methods, to better predict the most relevant educational material for each employee or group of employees, based on their demographics, organisational and situational characteristics. Some of these characteristics are likely to evolve over time. Therefore, the learning mechanisms should be able to dynamically process this data. Such information is available in most organisations deploying phishing simulations, since they run the studies regularly, providing measures of vulnerability and capability to report over time.

Finally, we plan to incorporate psychological factors into the predictive model. Although high predictive accuracy can be achieved using behavioural features, we posit that another source of valuable information is the recipient's psychological model. For example, psychological factors, such as locus of control, impulsivity, and perceived threat severity have the potential to influence the decision-making processes and behaviour of the email recipients. In the future, we will construct psychological models of the employees and incorporate them into predictive models.

References

1. Nurul Akbar. 2014. Analysing persuasion principles in phishing emails. University of Twente.
2. Nurcan Alkış. (12) The impact of individual differences on influence strategies. *ResearchGate*. Retrieved February 16, 2018 from

https://www.researchgate.net/publication/282720170_The_impact_of_individual_differences_on_influence_strategies

3. Brandon Atkins and Wilson Huang. 2013. A Study of Social Engineering in Online Frauds. *Open Journal of Social Sciences* 01, 03: 23–32. https://doi.org/10.4236/jss.2013.13004
4. Jan-Willem Bullee, Lorena Montoya, Marianne Junger, and Pieter Hartel. 2017. Spear phishing in organisations explained. *Information and Computer Security* 25, 5: 593–613. https://doi.org/10.1108/ICS-03-2017-0009
5. Marcus Butavicius, Kathryn Parsons, Malcolm Pattinson, and Agata McCormac. 2016. Breaching the Human Firewall: Social engineering in Phishing and Spear-Phishing Emails. In *Australasian Conference on Information Systems 2015*. Retrieved from http://arxiv.org/abs/1606.00887
6. Robert B. Cialdini. 2001. *Influence: Science And Practice*. Allyn And Bacon, Boston, MA.
7. Dan Conway, Ronnie Taib, Mitch Harris, Kun Yu, Shlomo Berkovsky, and Fang Chen. 2017. A qualitative investigation of bank employee experiences of information security and phishing. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, 115-129.
8. K. Coronges, R. Dodge, C. Mukina, Z. Radwick, J. Shevchik, and E. Rovira. 2012. The Influences of Social Networks on Phishing Vulnerability. In *2012 45th Hawaii International Conference on System Sciences*, 2366–2373. https://doi.org/10.1109/HICSS.2012.657
9. Julie S. Downs, Mandy Holbrook, and Lorrie Faith Cranor. 2007. Behavioral Response to Phishing Risk. In *Proceedings of the Anti-phishing Working Groups 2nd Annual eCrime Researchers Summit* (eCrime '07), 37–44. https://doi.org/10.1145/1299015.1299019
10. A. Ferreira and G. Lenzini. 2015. An analysis of social engineering principles in effective phishing. In *2015 Workshop on Socio-Technical Aspects in Security and Trust*, 9–16. https://doi.org/10.1109/STAST.2015.10
11. Ana Ferreira, Lynne Coventry, and Gabriele Lenzini. 2015. Principles of Persuasion in Social Engineering and Their Use in Phishing. In *Human Aspects of Information Security, Privacy, and Trust*, 36–47.
12. Edwin Donald Frauenstein and Rossouw von Solms. 2009. Phishing: How an Organization can Protect Itself. In *Information Security South Africa Conference 2009 (ISSA2009)*. Retrieved February 16, 2018 from https://www.researchgate.net/publication/220803149_Phishing_How_an_Organization_can_Protect_Itself
13. Tom N. Jagatic, Nathaniel A. Johnson, Markus Jakobsson, and Filippo Menczer. 2007. Social Phishing. *Commun. ACM* 50, 10: 94–100. https://doi.org/10.1145/1290958.1290968
14. Katharina Krombholz, Heidelinde Hobel, Markus Huber, and Edgar Weippl. 2015. Advanced Social Engineering Attacks. *J. Inf. Secur. Appl.* 22, C: 113–122. https://doi.org/10.1016/j.jisa.2014.09.005
15. Elmer EH Lastdrager. 2014. Achieving a consensual definition of phishing based on a systematic review of the literature. *Crime Science* 3, 1: 9. https://doi.org/10.1186/s40163-014-0009-y
16. J. G. Mohebzada, A. E. Zarka, A. H. Bhojani, and A. Darwish. 2012. Phishing in a university community: Two large scale phishing experiments. In *2012 International Conference on Innovations in Information Technology (IIT)*, 249–254. https://doi.org/10.1109/INNOVATIONS.2012.6207742
17. Nicole L. Muscanell, Rosanna E. Guadagno, and Shannon Murphy. 2014. Weapons of Influence Misused: A Social Influence Analysis of Why People Fall Prey to Internet Scams. *Social and Personality Psychology Compass* 8, 7: 388–396. https://doi.org/10.1111/spc3.12115
18. Charles Ohaya. 2006. Managing Phishing Threats in an Organization. In *Proceedings of the 3rd Annual Conference on Information Security Curriculum Development* (InfoSecCD '06), 159–161. https://doi.org/10.1145/1231047.1231083

19. Daniela Oliveira, Harold Rocha, Huizi Yang, Donovan Ellis, Sandeep Dommaraju, Melis Muradoglu, Devon Weir, Adam Soliman, Tian Lin, and Natalie Ebner. 2017. Dissecting Spear Phishing Emails for Older vs Young Adults: On the Interplay of Weapons of Influence and Life Domains in Predicting Susceptibility to Phishing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17), 6412–6424. https://doi.org/10.1145/3025453.3025831
20. Kathryn Parsons, Agata McCormac, Malcolm Pattinson, Marcus Butavicius, and Cate Jerram. 2015. The design of phishing studies: Challenges for researchers. *Computers & Security* 52: 194–206. https://doi.org/10.1016/j.cose.2015.02.008
21. Dawn M. Sarno, Joanna E. Lewis, Corey J. Bohil, Mindy K. Shoss, and Mark B. Neider. 2017. Who are Phishers luring?: A Demographic Analysis of Those Susceptible to Fake Emails. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61, 1: 1735–1739. https://doi.org/10.1177/1541931213601915
22. Steve Sheng, Mandy Holbrook, Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Julie Downs. 2010. Who Falls for Phish?: A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10), 373–382. https://doi.org/10.1145/1753326.1753383
23. Alex Tsow and Markus Jakobsson. 2007. *Deceit and Deception: A Large User Study of Phishing*. Indiana University, School of Informatics, Computing and Engineering, Bloomington. Retrieved February 16, 2018 from https://www.cs.indiana.edu/cgi-bin/techreports/TRNNN.cgi?trnum=TR649
24. Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D. & Wetherell, M. S. (1987). Rediscovering the social group: A self-categorization theory. Oxford: Blackwell
25. Arun Vishwanath, Brynne Harrison, and Yu Jie Ng. 2016. Suspicion, Cognition, and Automaticity Model of Phishing Susceptibility. *Communication Research*: 0093650215627483. https://doi.org/10.1177/0093650215627483
26. Emma J. Williams, Joanne Hinds, and Adam N. Joinson. 2018. Exploring susceptibility to phishing in the workplace. *International Journal of Human-Computer Studies* 120: 1–13. https://doi.org/10.1016/j.ijhcs.2018.06.004
27. Olga A. Zielinska, Allaire K. Welk, Christopher B. Mayhorn, and Emerson Murphy-Hill. 2016. A Temporal Analysis of Persuasion Principles in Phishing Emails. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 60, 1: 765–769. https://doi.org/10.1177/1541931213601175
28. *2017 Cost of Cyber Crime Study*. Accenture. Retrieved from https://www.accenture.com/t20170926T072837Z__w__/us-en/_acnmedia/PDF-61/Accenture-2017-CostCyberCrimeStudy.pdf